APRIL/MAY 2025

CECA63A/CECS63A/FECA63A/FECS63A — BIG DATA ANALYTICS

Time : Three hours Maximum : 75 marks

SECTION A — (10 × 2 = 20 marks)

Answer ALL questions.

1. Define Big Data.

2. List the main characteristics of Big Data.

3. Differentiate Data Stream Mining and Traditional Data Mining.

4. How Data are sampled in a stream?

5. Write down the two major functionalities of Hadoop YARN.

6. Give the properties of Stream Data.

7. Why is HDFS preferred to RDBMS?

8. Define HDFS.

Thiruvalluvar Arts & Science College T.V.Malai

9. Mention any two uses of Hive.

10. List the components of Hadoop Eco system.

SECTION B — (5 × 5 = 25 marks)

Answer ALL questions.

11. (a) What is the need for Big Data Analysis?

Or

(b) Give an overview on Big Data Evolution.

12. (a) Classify the approaches used to estimate the moments in a stream.

Or

(b) What is RTAP? Why it is important in timely Decision Making?

13. (a) How Hadoop Streaming is suited with text processing? Explain.

Or

(b) Explain the building blocks of Hadoop with neat sketch.

Sriammuga Industries Arts & Science College T.V.Malai

14. (a) How do you analyse the data in Hadoop?

Or

(b) What are the configuration parameters in a 'Map reduce' program?

15. (a) Explain two execution types or modes in PIG.

Or

(b) How will you query the data in Hive?

SECTION C — (3 × 10 = 30 marks)

Answer any THREE questions.

16. Explain the characteristics, features and the structure of big data in detail.

17. How is Sentiment Analysis playing a major role in Data Mining?

18. What is Hadoop? Explain its components.

19. Discuss the various system roles in an HDFS components or development.

20. Describe IBM info sphere Big insights and streams concept.

———————